2004-05-03

Time Frame and its Impact on Commodity Trading Advisor Performance

Nordia D. Thomas
Worcester Polytechnic Institute

TIME FRAME AND ITS IMPACT ON COMMODITY TRADING ADVISOR PERFORMANCE

by

Nordia D. Thomas

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

in

Financial Mathematics

by

____

May 2004

APPROVED:

____

Prof. Kathryn Wilkens, Thesis Co-advisor

____

Prof. Bogdan Doytchinov, Thesis Co-advisor

____

Prof. Bogdan Vernescu, Head of Department

Abstract

This thesis explores the idea that time frame is an important determinant of commodity trading advisor (CTA) performance. Results allow us to reject the hypothesis that short-term price movements may be due only to noise, thus CTAs will have the same performance regardless of time frame. Using several performance measures and multi-factor models we find instead that CTAs who focus on short-term price changes are better positioned to benefit from advances in financial information processing and trade execution technology.

Acknowledgements

I wish to extend my deepest thanks to an anonymous referee for giving me insight into the world of commodity trading advisors, for helping me understand the terminology, putting aside time to help with the fine tuning of the survey which was one of the most important products of this research, and for making my talks with various commodity trading advisors go as smoothly as possible. I would like to thank Prof. Wilkens for helping me come up with a new thesis topic that involved futures when I informed her I was not interested in doing a thesis on VaR. I would also like to thank Prof. Doytchinov who came onboard the project later on. I would also like to thank Prof. Vermes for his support throughout my two years in the Financial Mathematics program. I would also like to thank the various commodity trading advisors who took time out of their busy schedules to fill out and submit the survey. Thank you all!

Worcester, March 2004
Nordia Thomas

To Hyter, Youlonda, David, Neil, and Jennifer

1 Introduction

Trading in futures in the United States is regulated by the Commodity Futures Trading Commission (CFTC). The CFTC "protects market participants against manipulation, abusive trade practices and fraud... [and] enables the markets to serve better their important functions in the nation's economy—providing a mechanism for price discovery and a means of offsetting price risk" [CTC website]. Professionals, who primarily focus their trading in futures, and in some cases options on futures, are known as commodity trading advisors (CTAs). Some run separately managed accounts where they act as consultants, while others run their own fund(s). All CTAs are required to register with the CFTC.

CTAs collectively comprise the alternative investment area called managed futures. This area is made up of different strategies and sub-strategies. Managed futures are a vital component in a portfolio because they have little or no correlation with stocks. Because of this, adding futures to a portfolio of stocks, or even a portfolio of stocks only can increase the diversification of that portfolio.

In this thesis, we test the hypothesis that short-term price movements may be due to only noise and CTAs have equal performance regardless of time frame. An alternative hypothesis is that CTAs who focus on short-term price changes are better positioned to benefit from ongoing advances in financial information processing and trade execution technology. These CTAs use models with greater sensitivity to changes in underlying asset prices and volatility and thus enter and exit positions more quickly than their longer-term, less sensitive counterparts.

CTAs either explicitly or implicitly believe that market prices do not move randomly seek to profit from price inefficiencies. There are various types of CTAs. Most of these managers are trend-followers. Most trend followers, in turn, are systematic traders, i.e., they develop models to identify trends created by inefficient markets and to generate entry and exit signals to capture those trends. The time frames in which they analyze and attempt to capture trends vary widely, from minutes to hours to months a year or more. Some CTAs specialize in one time frame in the belief that market inefficiencies occur more frequently within it, whereas others run a multi-time frame set of models, believing that inefficiencies occur across different time frames and that time-frame diversification is vital.

Virtually all studies on CTAs have used monthly data, and because of the proprietary nature of CTA models so shall we. However, we use a unique data set based on monthly CTA return data combined with survey data. The survey questions are carefully designed to elicit information on the price change sensitivity of CTA trading models in such a way that we can classify managers and then evaluate their performance to identify the relationship between model sensitivity and returns. We describe the various criteria used to identify different degrees of sensitivity, such as the percentage change in the market that triggers entry and/or exit signals in a trader's model. We also analyze Sharpe ratio, Sortino ratio, Calmar ratio, and two equilibrium-based factor models to determine trading success based on use of information flow.

We are interested in this issue because alongside the growth in the managed futures industry, investors have shown increasing interest in short-term CTAs over the last several years. However, current and prospective investors in CTAs often lack sufficient information or a framework within which to analyze the time frame sensitivity of CTAs' models. Moreover, we believe that time sensitivity provides one more factor investors can use in the hope of achieving diversification among the managers they invest in.

The thesis is organized as follows. In Chapter 2 we present a brief literature review. Chapter 3 presents theoretical concepts and various methodologies used to evaluate CTA performance; it also outlines the different sources of data and survey details. Chapter 4 describes empirical results and analysis, and in Chapter 5 we conclude and make suggestions for further study.

2 Literature Review

2.1 Alternative vs. Traditional Investments

We present research that is based on the premise that time frame is an important determinant of CTA performance. The conventional way traders refer to time frame is how long on average a trader holds a position before reversing it, exiting it, or going into neutral mode. We however think of time frame in terms of the sensitivity of a trading model to market changes; for example, we look at what degree of change in price or volatility will cause an entry or exit in the model. For this research we refer to 0-6 days as *short-term*, 7-30 days as *intermediate-term*, and greater than 30 days as *long-term* time frames.

While there has been some proprietary institutional research that suggests that time frame is the most important determinant of CTA performance, all things being equal, there is little evidence reported in the literature. The research presented here is in some ways similar to the 1986 landmark study by Brinson, Hood, and Beebower (hereafter referred to as BHB). BHB showed that for 91 large U.S. pension plans where data was analyzed over the period 1974-1983, on average 93.6 percent of the variation in plan returns was due to asset allocation. In 1991 Brinson, Singer, and Beebower conducted a follow-up study. The updated study used 1977-1987 data for 82 pension plans and showed that asset allocation determined 91.5 percent of plan return performance.

Here, however, we pose the performance question in the context of managed futures. We know that for pension plans and traditional mutual funds with long-only positions and relatively passive investment strategies, that asset allocation is most responsible for return variation. We want to know for CTAs, who frequently shift allocations among markets and sectors using futures and who take long and short positions extensively, and often with rapid turnover, what impacts performance.

In 1992, Sharpe produced an asset allocation model for mutual funds that utilized asset class factors. Sharpe used only eight "major" asset classes to replicate mutual fund performance. However, because of the differences between the trading strategies of mutual fund managers and those of hedge fund managers/CTAs, Sharpe's asset class factor model cannot be used for hedge funds/CTA pools without modification. In a 1997 study Fung and Hsieh discuss an extension of Sharpe's model. They assume that the key determinants of hedge fund/CTA pools performance were location, trading strategy, and leverage. Their asset class factor model extended Sharpe's by replacing the original eight asset classes with more global asset classes, adding high yield corporate bonds as a new asset class, and including three dynamic trading strategies: Systems/Trend Following, Systems/Opportunistic, and Global/Macro. Unfortunately, Fung and Hsieh met with limited success. Unlike Sharpe's model which showed high correlation with nearly all mutual funds, their model is highly correlated with only 40 percent of hedge funds/CTA pools. The main problem with their model is that you cannot use monthly data in a regression to figure out positions in asset classes that change more frequently than monthly. We are undertaking this study because of the inability of this model to ascertain key determinants of hedge funds/CTA pools.

2.2 Market Efficiency

CTAs are often classified as a type of hedge fund manager. They apply active portfolio management techniques, taking long and short positions, and generally trade in a diverse set of markets in an attempt to capture market inefficiencies. The classic definition of market inefficiency implies that prices instantaneously and fully reflect all available information [Fama 1991]. Most academics and practitioners agree that markets are not fully efficient; however they have different views on the degree of efficiency.

Most theorists believe that there are three degrees, or forms, of market efficiency. *Weak form* efficiency states that past performance is not an indicator of future performance, that is, all information from past prices is reflected in current prices. *Semi-strong form* efficiency states that investors will not perform better than the market unless they have information other than publicly available information. *Strong form* efficiency says that even if an investor has all information including private information he or she still will not perform better than the market. If an investor dud however manage to generate above normal returns this is due purely to chance.

However there is empirical evidence of systematic patterns in asset returns that challenge the theory that the market is efficient. Among these anomalies are calendar effects such as the *weekend effect*, the *holiday effect*, and the *January effect*. In the weekend effect, formerly known as the Monday effect, researchers noted that returns on Mondays were more likely to be negative than returns on other days of the week. In the holiday effect mean stock returns are high on the day prior to the start of the holiday. Of all these calendar effects, the January effect is the most studied, and researchers have found that returns in January tend to be much higher than those in the rest of the year [Arsad and Coutts 1993; see Wilkens (2000) for a review of these and other patterns in return and volatility in cash and futures markets].

3 Data and Methodology

3.1 Risk Measures

In simple terms, risk is the exposure to unexpected change. For investors it could be considered the exposure to unexpected returns. But risk is hard to define and even more difficult to measure [Jaeger 2000]. Below are several measures that are used in the industry to estimate risk.

3.1.1 Standard Deviation

Standard deviation (σ) is one of the most common measures of historical volatility, and is often used as a proxy for risk. It is the square root of variance, which measures the average squared difference between actual returns and average returns.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(R_i - \bar{R})^2}{n-1}}$$

where σ is the standard deviation

R_i is the actual return at time i

$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i$ is the average return

n is the length of the time period

As a risk measure standard deviation fails to distinguish between risk and uncertainty. In addition it penalizes upside volatility, though investors are averse to poor performance not overperformance [Markowitz 1992].

3.1.2 Downside Deviation

Downside deviation σ_{DD} is the standard deviation of below target semivariance, which measures the dispersion of returns about a minimum acceptable return (MAR). Unlike standard deviation, downside deviation differentiates between risk and uncertainty, penalizes only for underperformance, and does not imply a symmetric, normal return distribution.

$$\sigma_{DD} = \sqrt{\frac{\sum_{i=1}^{T}(R_i - MAR)^2}{T}}$$

where T is the number of periods where $R_i \le MAR$, and $T \le n$.

3.1.3 Drawdown

Drawdown (DD) is the percentage loss in the program's asset value from a peak to a subsequent trough. The maximum drawdown is the maximum cumulative loss over the entire investment record (or for a particular period), and it is often used as a "key measure of track record quality and strategy 'riskiness'" [Harding, Nakou, Nejjar 2002]. For a particular time period we look at the net asset value (NAV) we determine the peaks or high-water points. The drawdown is then computed as the percentage change between the current NAV and the high-water mark.

$$\mathrm{DD}_t = \frac{\mathrm{NAV}_t - \mathrm{Highwater}}{\mathrm{Highwater}}$$

3.1.4 Average True Range

Average true range is an indicator of volatility. It was introduced by J. Welles Wilder in his book *New Concepts in Technical Trading Systems*. True range is the maximum of the following:

- the difference between the current high and the current low
- the absolute value of the difference between the current high and the previous close
- the absolute value of the difference between the current low and the previous close

The average true range is the moving average of the true ranges over a 14 day period. Although we do not go on to compute average true range in this research, it is not because we are using monthly instead of daily data. In fact, average true range is used for underlying data, so CTA data frequency is not an issue. True range and implied volatility are not commonly used on CTA data, instead CTAs use these risk measures to make their trades.

3.1.5 Implied Volatility and VIX

If we calculate the value of an option using the Black-Scholes model it would differ from the market price of the option.[1] The implied volatility is the value of

---

[1] Non-dividend paying European options are priced using the Black-Scholes model as
$c = S_0 N(d_1) - Ke^{-rT} N(d_2)$ and $p = Ke^{-rT} N(-d_2) - S_0 N(d_1)$
where c is the European call price
p is the European put price

$$d_1 = \frac{\ln(S_0/K) + \left(r + \sigma^2/2\right)T}{\sigma\sqrt{T}}$$

volatility that would equate the current option price with the Black-Scholes price [Hull 2002]. VIX is the ticker symbol for the Chicago Board Options Exchange (CBOE) Volatility Index. Prior to September 2003 VIX was calculated by taking a weighted average of the implied volatility from eight calls and puts from the Standard and Poor's 100 Index (OEX). The options were weighted by their time to maturity and how far in or out of the money they are, leading resulting in a hypothetical at-the-money option that would expire in 30 days. Since the introduction of VIX in 1993 there have been many changes in the way both practitioners and theoreticians view volatility [CBOE 2003]. Because of this a new way to calculate stock market volatility was introduced. This new index has ticker VIX while the old index is now referred to by the ticker symbol VXO.

The new VIX differs from the old in many ways[2]. The two most important are:

- Calculation of implied volatility is not based on the Black-Scholes pricing model, does not depend on any model, and involves options with a wider range of strike prices than does VXO [CBOE 2003].
- VXO used OEX options for its calculations because they were the most traded and most liquid index options on the CBOE. However the new VIX uses options on the Standard and Poor's 500 Index (SPX) which is highly correlated with the OEX, is a common benchmark for U.S. stocks, and is

---

$d_2 = d_1 - \sigma\sqrt{T}$, and

$N(\cdot) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$, is the standard cumulative normal distribution

[2] For more details about the new VIX you can read the CBOE VIX white paper. It can be found at http://www.cboe.com/micro/vix/vixwhite.pdf.

comprised of many of the most traded assets for which options are created [CBOE 2003].

VIX is useful because it has an inverse relationship with the underlying stock market index, this is illustrated in Figure 1 below.

Figure 1: Plot of daily S&P 500 and VIX values over the period January 2002 -- December 2003.

VIX is considered to be the 'investor's fear gauge', this is so because "VIX is based on real-time option prices, which reflect investors' consensus view of future expected stock market volatility. Historically, during periods of financial stress, which are often accompanied by steep market declines, option prices - and VIX - tend to rise. The greater the fear, the higher the VIX level. As investor fear subsides, option prices tend to decline, which in turn causes VIX to decline" [CBOE 2003].

3.1.6 Lookback straddle

A straddle involves the purchase of an equal number of call and put options with the same terms, i.e., with the same strike price and expiration date, at the same time. A straddle allows an investor to make a profit from expected volatility in the market even if he or she is uncertain in which direction the market will move.

When Fung and Hsieh extended the Sharpe (1992) asset allocation model they realized that dynamic trading strategies better explained hedge fund/CTA pools returns than did traditional static strategies. In the course of their research they "showed that the returns from trend-following strategies can be replicated by a dynamically managed option-based strategy" [Fung and Hsieh 2002]. The option that best models these strategies is a lookback straddle. A lookback straddle involves the purchase of an equal number of lookback calls and puts with the same strike price and expiration date. A lookback call allows its owner to purchase an asset at its lowest price, while a lookback put allows its owner to sell an asset at its highest price within some given period. The payoff from the lookback straddle is therefore pays the difference between the maximum and the minimum value of the asset in question [Hull; Kolb]. That is,

$$
\begin{aligned}
LBC &= \max(0, S_T - \min(S_t, S_{t+1}, \ldots, S_T)) \\
LBP &= \max(0, \max(S_t, S_{t+1}, \ldots, S_T) - S_T)
\end{aligned}
$$

where LBC is the payoff for a lookback call, LBP is the payoff for a lookback put, and S_t is the price of the underlying asset at time t.

3.2 Performance Measures

There are many measures to evaluate how well a CTA is performing. Among them are the Sharpe ratio, Sortino ratio, Calmar ratio, and alphas from single- and multi-factor models.

3.2.1 Sharpe Ratio

The Sharpe ratio for a given CTA program is defined as the ratio of excess return to standard deviation. That is

$$\text{Sharpe Ratio} = \frac{\bar{R}-\bar{R}_f}{\sigma}$$

where $\overline{R}$ is the average return

$\overline{R}_f$ is the monthly return of the 90-day Treasury bill

σ is the standard deviation of the returns

The Sharpe ratio was devised by William Sharpe in 1966 for use in deciding how desirable an investment was. Since then it has become very widely used because it is easy to calculate and what is represents is easy for investors to understand. However, the Sharpe ratio has drawbacks which stem from its dependence on standard deviation as a measure of risk. The standard deviation takes into account the distance of each return from the mean, and as such can be overly influenced by outliers in the data set. Because of this, large positive returns are treated no differently from large negative returns, causing penalization of those CTA programs with above average return. This is direct contrast to the fact that investors are loss averse and view large positive returns and large negative returns very differently.

3.2.2 Sortino Ratio

To improve upon the Sharpe ratio the Sortino ratio was designed. It replaces the benchmark return with the minimum acceptable return (*MAR*) and replaces the standard deviation with the downside deviation (σ_{DD}):

$$\text{Sortino Ratio} = \frac{\bar{R} - \text{MAR}}{\sigma_{DD}}$$

The problem with the Sortino ratio is that while it only penalizes for negative returns, each program can have different Sortino ratio depending on the value of *MAR*.

3.2.3 Calmar ratio

Another risk/return ratio is the Calmar ratio. It is defined as

$$\text{Calmar Ratio} = \frac{\text{ROR}}{|\max(\text{DD})|}$$

where *ROR* is the compounded annualized rate of return. The problems with the Calmar ratio stem from its dependence on drawdown as a statistical measure of risk. If a manager reports daily data he is expected to have more drawdowns than a manager who reports weekly data, and so on. In addition the longer a program is active, the larger its maximum drawdown is likely to be. If there is no correction for these two issues the Calmar ratio is likely to penalize for this.

3.2.4 Single-factor Models

The last two performance measures we look at are two linear unconditional factor models. The first is a single factor model while the second is a multi-factor model. The single factor model is expressed as

$$E[R_i] - R_f = \alpha_i + \beta_i \left(E[R_m] - R_f\right), \quad 1 \le i \le ???$$

where R_i is the total random return on the i th CTA program, R_f is the one period (monthly?) total return on the riskless asset (we use the Bloomberg Generic Treasury 3-month bill rate), R_m represents the total return on the benchmark (we use the Barclay CTA Index and the S&P Managed Futures Index), α_i is the abnormal performance of the i th CTA program, and β_i is a measure of the volatility of the i th CTA program relative to the benchmark.

The single factor model we use is based on the Capital Asset Pricing Model (CAPM). CAPM describes the relationship between risk and expected return for ???. Its theoretical form is

$$E[R_i] - R_f = \beta_i \left(E[R_m] - R_f\right)$$

The empirical form is

$$E[R_i] - R_f = \alpha_i + \beta_i \left(E[R_m] - R_f\right), \quad \alpha_i \neq 0$$

where α_i is a measure of the excess return over what CAPM predicts.

CAPM was developed by William Sharpe, John Litner, and Jan Mossin. It provides a simple way to think of risk and reward, as it quantifies the tradeoff between risk and reward by providing a "precise prediction of the relationship between an asset's risk and its expected return." CAPM is based on the idea that:

- investors are risk averse and have a one-period investment horizon,
- investors are mean-variance optimizers i.e. they use mean and standard deviation as criteria to decide where to invest their money,
- mean and standard deviation data is available,

- assets are perfectly divisible and investors are price-takers, i.e., there are many investors and they cannot influence the market,
- there is unlimited borrowing and lending at the risk free rate,
- all assets are marketable,
- there are complete markets,
- investors have the same information so they have the same expectations re mean, variance, covariance, etc., and
- there are zero taxes and no transaction costs.

We do not use a single-factor model in this research. While there are some CTAs who invest in only one market sector, most CTAs trade in several markets. Our decision not to use a single-factor model is okay because determining what could be an appropriate market proxy would be quite difficult.

3.2.5 Multi-factor Models

The multi-factor model we are using grew out of the fact that the CAPM β does not completely explain the expected returns for the asset in question. The model is based on the Arbitrage Pricing Theory (APT) introduced in 1976 by Sheldon A. Ross. The APT is based on the idea that there are no arbitrage opportunities in the marketplace. In everyday terms this is the assumption that "there is no such thing as a free lunch." What this means in financial terms is that one cannot make a profit without incurring some risk.

The APT is intended to overcome some of the limitations of CAPM. Like CAPM, the APT assumes that markets are frictionless with no taxes and zero transaction costs. Unlike CAPM however, the APT does not require us to identify the market portfolio and gives only an "approximate relationship between an asset's risk and its expected return." It is based on the premise that an asset's expected return is based on different factors, that is,

$$E\left[R_i\right] - R_f = \alpha_i + \beta_{i1}F_1 + \beta_{i2}F_2 + \cdots + \beta_{in}F_n$$

where β_{ij} is a measure of the sensitivity of the returns of the i th asset to the j th factor, and F_j is the j th factor.

The multi-factor model is expressed as

$$E[R_i] - R_f = \alpha_i + \sum_{j=1}^{m} \beta_{ij}\left(E[R_j] - R_f\right),$$

where R_i is the total random return on the i th CTA program, R_f is the one period (monthly?) total return on the riskless asset, R_j represents the total return on the j th benchmark, α_i is the abnormal performance of the i th CTA program, and β_{ij} is a measure of the volatility of the i th CTA program relative to the j th benchmark.

We assess R_j for different time frames. CTAs programs fall into anyone of three time frames: short-term, intermediate-term, and long-term. For the multi-factor model the additional factors are the S&P 500 Total Return Index for equities; the Lehman U.S. Aggregate for fixed income; the GSCI Energy, GSCI Metal, and GSCI Non-energy indices for commodities, and the NEXN for foreign exchange.

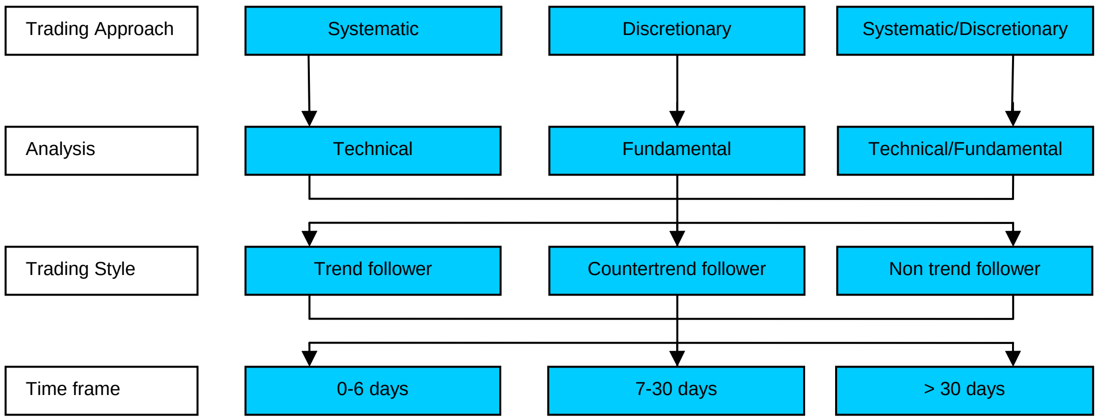
Using the multi-factor model we also try to determine the relation between CTA returns and volatility proxies for our basic market sectors. The volatility proxies all represent the return on a straddle using options with one month to maturity. The series is calculated by the CISDM using historical implied volatilities and a risk free rate of 5%. The indices used are the S&P 100 Index (OEX), the future on the 5-year (Treasury) note, the GSCI Euro-currency futures contract, and the GSCI futures contract for equities, fixed income, foreign exchange, and commodities respectively.

3.3 Data and Survey Details

CTAs have three trading approaches: they are either systematic traders, discretionary traders, or employ some combination of the two approaches [Lungarella 2002]. Systematic traders develop computer-based mathematical models to analyze historical prices to identify trends created by inefficient markets that can be used to forecast market movements and generate entry and exit signals [Ghaleb-Harter 2004]. Discretionary traders in the other hand rely on fundamental analysis of market conditions to determine their trading strategies.

There are three overall trading styles: trend following, counter trend following, and non-trend following. Most CTAs are trend-followers. An overview of the various investment approaches of CTAs is shown in Figure 2 below:

Figure 2: Overview of the various investment strategies used by CTAs. (adapted from [Lungarellla 2002]).



Investors in the managed futures industry need a way to evaluate the performance of managers. To this end we implemented a survey aimed at determining the impact of time frame and related style factors on CTA performance.

Since most of the models used by systematic managers are proprietary, information on these models is kept close to the vest. Consequently, we carefully composed the questions in the survey, with suggestions from industry practitioners, to learn as much as possible. Some of the questions in the survey are typical due diligence questions, while the others are specifically formulated to extract as much information as possible about the price and/or volatility sensitivity of the managers' models. The complete survey can be found in Appendix A.

We obtained the contact information for the CTA firms listed in the International Traders Research, Inc. (ITR) database. We then contacted them by email, sent follow-up emails, and requested that they complete the online survey. As an incentive we promised a copy of the final research write-up to each CTA that completed the survey. Of the 202 CTAs listed, only 21 completed the survey, a response rate of 10.4 percent.

Such a low response by the CTAs was unexpected even after we had taken into account the fact that the survey was aimed towards systematic traders whereas the database contained the names of both systematic and discretionary traders. After waiting an additional week to see if more results would come in and making several calls to CTAs, we decided to reduce the number of questions we asked. Since we felt that firms were concerned about the proprietary nature of their data we asked instead that they respond to the following three questions for each of their programs:

1. What market sectors do you trade?
   Equities Percent of trading ___%
   Fixed Income Percent of trading ___%
   Foreign Exchange Percent of trading ___%
   Commodities Percent of trading ___%

2. What is the frequency of the data you use to generate trading signals? (e.g. 5-min, daily, weekly) _____________

3. What is your average holding period?
   Overall _____________
   For winning trades _____________
   For losing trades _____________

Once this follow-up email was sent we received replies from an additional 22 CTA programs. Of these 43 CTA programs only 30 had returns for the period January 2002 to December 2003 *and* had not changed their trading strategy within that period such that it led to a change in their average holding period. We were able to access monthly returns data for these eligible CTAs in the ITR CTA database. We chose this sampling frequency because of the unavailability of daily returns data from a large enough number of CTAs to make results statistically significant and representative of the larger CTA population. Summary statistics of the monthly returns for the eligible CTAs are shown in Table 1.

Table 1: Summary Statistics for the 30 eligible CTA programs

Program	Mean ROR	Minimum ROR	Maximum ROR	Standard Deviation	Downside Deviation	Maximum Drawdown
A	4.28%	-13.55%	32.78%	10.15%	3.60%	-20.24%
B	0.46%	-13.83%	10.59%	5.08%	4.08%	-14.40%
C	1.70%	-9.73%	13.90%	6.22%	2.99%	-5.15%
D	2.03%	-15.59%	12.99%	7.30%	4.68%	-29.73%
E	0.86%	-3.49%	5.20%	2.11%	1.05%	-8.93%
F	0.85%	-2.90%	5.30%	2.08%	0.95%	-16.67%
G	0.29%	-8.07%	7.74%	4.99%	2.26%	-85.55%
H	1.10%	-3.09%	3.31%	1.70%	0.36%	-28.14%
I	0.28%	-3.90%	3.47%	1.89%	1.27%	-23.08%
J	0.70%	-5.62%	6.62%	3.38%	1.75%	-27.01%
K	1.71%	-13.38%	39.56%	12.90%	3.92%	-34.49%
L	0.01%	-1.72%	3.76%	1.44%	0.54%	-92.04%
M	-0.29%	-2.95%	4.39%	1.86%	0.96%	-66.00%
N	1.27%	-7.69%	14.90%	4.87%	3.01%	-27.67%
O	0.78%	-6.84%	10.68%	4.78%	2.48%	-18.24%
P	0.26%	-5.31%	7.01%	3.93%	1.45%	-9.91%
Q	0.34%	-5.31%	8.66%	4.20%	1.55%	-56.43%
R	2.65%	-20.94%	47.02%	14.62%	6.05%	-30.74%
S	3.41%	-11.73%	26.86%	11.11%	3.94%	-27.31%
T	1.21%	-12.49%	10.20%	5.80%	3.50%	
U	0.57%	-3.54%	6.19%	2.10%	1.12%	-2.47%
V	-0.48%	-8.20%	7.58%	4.62%	3.12%	
W	0.52%	-5.81%	4.61%	2.74%	1.65%	-2.90%
X	8.62%	-23.82%	45.32%	19.01%	5.53%	-69.09%
Y	1.56%	-13.78%	16.82%	7.35%	3.83%	-7.87%
Z	3.91%	-17.35%	31.93%	14.39%	5.19%	-17.21%
AA	0.77%	-9.29%	16.56%	6.38%	2.92%	-42.19%
AB	2.51%	-7.41%	12.69%	6.06%	2.54%	-36.70%
AC	0.78%	-10.06%	13.86%	7.37%	2.75%	-50.00%
AD	2.33%	-1.06%	7.38%	2.38%	0.14%	-11.34%

From the table we see that the average monthly returns for these CTAs is positive, with the exception of Programs M and V where the average return is –0.29% and –0.48% respectively.

3.4 Data Problems: Biases and other Effects

When we determine the performance of CTAs using returns data we have to be careful with our inference. This is because CTA returns data is subject to several measurement biases caused by the way the data is collected by the ITR database, the way the data is collected for this particular research, as well as the way the managed futures industry is organized. Survivorship bias comes about when only data from surviving CTA programs is used to calculate performance [see New 2001; Schneeweis, Spurgin and McCarthy 1996].

This can result in an upward or downward bias. For the sample period it results in upward bias because some non-surviving programs which were omitted in the analysis are probably defunct because they had poorer performance than the surviving CTA programs. It results in downward bias because successful programs which have achieved full capacity, and are no longer trying to attract investors may choose to stop reporting returns even if they are performing well. Because the ITR CTA database does not have data for the entire universe of CTAs, nor does it have data for failed CTA programs, all our performance estimates are affected by survivorship bias.

New [2001] talks about "self-selection bias" for hedge fund managers and his arguments hold true for managed futures managers. Like hedge fund managers, CTAs voluntarily report to the ITR CTA database and other CTA databases. Managers whose programs are doing well might seek inclusion, whereas managers who are doing poorly might choose to avoid reporting their returns. This selection of self-reporting bias can be such that the CTAs in the database are not representative of the population of CTAs.

CTA programs usually go through an incubation period, so prior to being added to a database they will already have returns data. If incubation is successful the programs are marketed even further and if unsuccessful the programs are liquidated and their results are not included in the database in question. When the program is added to the database the past returns are 'backfilled' creating an instant history [Fung and Hsieh 2000]. This results in an upward bias in the performance of the programs.

Additional bias known as multi-period sampling bias is introduced because we restrict our analysis to a subsample of programs having returns data for the period January 2002 to December 2003. Since the managers did not complete the survey for each of their programs this induces additional selection bias since the managers who are willing to do the survey in the first place are more likely to do so for programs which are performing well. Indeed, it might be the case that a manager with several programs might choose the best program and use those statistics when filling out the online survey.

Since the CTAs in the sample suffer from these same biases and we are comparing them to each other, we are not concerned about these biases. We are however concerned about the fact that our sample size is small and we do not have longer return history for the CTA programs in our sample. It is possible that the results of this research do not extend to other time periods, and due to the small sample size the results may not be generalizable for the larger population of CTAs.

4 Empirical Results

We performed a multi-linear regression for the CTA programs against the market sector benchmarks, the volatility proxies, and the market sector benchmarks and volatility proxies respectively. The results are listed in the Tables 2—4 below.

Table 2: Regression Results for Multi-factor Model using Market Sector Benchmarks

ALPHA is the excess return, TRI is the S&P 500 Total Return Index, LUA is the Lehman U.S. Aggregate, GSCIE is the GSCI Energy, GSCIM is the GSCI Metals, GSCIN is the GSCI Non-energy, and NEXN is the JP Morgan Nominal Exchange Rate Index.

Dependent Variable	ALPHA	TRI	LUA	GSCIE	GSCIM	GSCIN	NEXN	Adj. R-Squared
Program A	0.0194	-0.3434	1.7947	0.1268	0.8336	-0.3251	-0.7272	-0.1026
	0.3945	-0.6206	0.9268	0.5020	1.2822	-0.3602	-0.3649	
Program B	-0.0344	-0.4786	0.5379	0.1254	0.4813	0.0696	-1.1691	0.3491
	-1.8443	-2.2805	0.7322	1.3096	1.9516	0.2035	-1.5467	
Program C	-0.0244	-0.2839	0.6729	0.3211	-0.1177	0.7302	-1.5658	0.4740
	-1.1673	-1.2071	0.8174	2.9912	-0.4258	1.9035	-1.8484	
Program D	-0.0298	-0.1766	1.1431	0.1534	0.3067	-0.0511	-2.1540	0.0698
	-0.9130	-0.4802	0.8881	0.9140	0.7098	-0.0852	-1.6263	
Program E	-0.0185	-0.0333	-0.1158	-0.0711	0.0240	-0.0464	-0.7492	-0.0382
	-1.8783	-0.3008	-0.2988	-1.4071	0.1846	-0.2571	-1.8791	
Program F	-0.0045	-0.1238	0.5829	-0.0658	0.2508	-0.0807	-0.1586	0.0426
	-0.4815	-1.1828	1.5912	-1.3773	2.0392	-0.4730	-0.4208	
Program G	0.0085	0.7112	-0.9009	0.1097	-0.3585	0.5329	0.8159	0.4335
	0.4767	3.5323	-1.2784	1.1941	-1.5152	1.6227	1.1251	
Program H	0.0067	0.0235	-0.0733	0.0015	-0.0818	0.1305	0.4408	-0.1342
	0.8444	0.2638	-0.2350	0.0364	-0.7811	0.8982	1.3738	
Program I	-0.0027	-0.0383	0.2777	-0.0174	0.0897	0.0189	0.3142	-0.2587
	-0.2740	-0.3460	0.7170	-0.3449	0.6901	0.1046	0.7882	
Program J	-0.0120	-0.3911	1.1288	0.1781	0.1777	0.2552	-0.8712	0.5778
	-0.9033	-2.6115	2.1533	2.6061	1.0100	1.0446	-1.6152	
Program K	-0.0162	-0.3978	1.0936	0.6226	-0.1936	-0.3173	-0.5911	0.1252
	-0.2901	-0.6333	0.4974	2.1718	-0.2623	-0.3098	-0.2613	
Program L	-0.0149	0.0379	0.2942	-0.0005	-0.0856	0.2119	-0.3223	0.0877
	-2.3465	0.5321	1.1788	-0.0157	-1.0221	1.8225	-1.2550	
Program M	-0.0245	-0.0281	-0.4851	0.0501	-0.0489	0.0528	-0.2299	-0.0965
	-2.8938	-0.2944	-1.4520	1.1489	-0.4357	0.3390	-0.6689	
Program N	-0.0335	-0.1189	1.0859	-0.0322	0.5035	-0.1269	-2.0857	0.3920
	-1.9143	-0.6044	1.5777	-0.3592	2.1791	-0.3958	-2.9449	
Program O	0.0024	-0.2697	2.1630	-0.0923	0.6990	-0.5099	-0.1071	0.2939
	0.1283	-1.2753	2.9228	-0.9561	2.8135	-1.4784	-0.1407	
Program P	0.0090	-0.2395	0.6117	0.0650	-0.1440	0.4612	0.8692	0.0863
	0.5228	-1.2356	0.9017	0.7353	-0.6325	1.4588	1.2451	

Dependent Variable	ALPHA	TRI	LUA	GSCIE	GSCIM	GSCIN	NEXN	Adj. R-Squared
Program Q	0.0147	-0.1924	0.6553	0.0870	-0.2405	0.6350	1.0409	0.1346
	0.8176	-0.9540	0.9282	0.9451	-1.0145	1.9300	1.4328	
Program R	-0.0330	-1.3646	2.8971	0.2531	1.2779	-0.0265	-2.4059	0.2218
	-0.5559	-2.0410	1.2380	0.8296	1.6266	-0.0243	-0.9991	
Program S	0.0747	0.0845	0.3356	-0.2216	1.3228	0.1435	2.6180	0.2693
	1.6756	0.1683	0.1911	-0.9680	2.2440	0.1754	1.4489	
Program T	-0.0156	-0.2415	0.5179	0.2853	-0.0973	0.6114	-0.7792	0.2496
	-0.6763	-0.9287	0.5692	2.4041	-0.3184	1.4419	-0.8322	
Program U	-0.0142	0.2320	-0.3523	0.0358	-0.2125	0.3472	-0.4825	0.2757
	-1.6571	2.4096	-1.0455	0.8138	-1.8788	2.2109	-1.3915	
Program V	-0.0316	-0.0001	-0.3774	0.0968	-0.4655	0.2796	-0.6466	-0.0448
	-1.4785	-0.0005	-0.4478	0.8804	-1.6451	0.7118	-0.7456	
Program W	-0.0067	0.1518	0.2094	0.1479	-0.3082	0.2462	-0.0831	0.0870
	-0.5483	1.1107	0.4379	2.3716	-1.9196	1.1047	-0.1689	
Program X	-0.0092	-0.5613	-0.0580	0.3600	1.8492	0.6972	-3.7748	0.1906
	-0.1159	-0.6293	-0.0186	0.8844	1.7645	0.4793	-1.1751	
Program Y	-0.0262	0.0765	0.0657	0.3903	-0.5207	0.3992	-1.4466	0.1396
	-0.8358	0.2165	0.0532	2.4207	-1.2545	0.6929	-1.1371	
Program Z	-0.0002	0.1814	1.2962	0.1425	-0.4586	0.3574	-2.0778	-0.2487
	-0.0030	0.2170	0.4429	0.3735	-0.4668	0.2620	-0.6899	
Program AA	-0.0364	-0.4490	-0.6798	0.1105	-0.2238	-0.2562	-0.9722	0.1138
	-1.3324	-1.4598	-0.6316	0.7871	-0.6194	-0.5107	-0.8777	
Program AB	-0.0106	-0.1410	1.7098	0.1840	0.2487	0.6643	-1.8244	0.4850
	-0.5293	-0.6229	2.1585	1.7816	0.9351	1.7997	-2.2383	
Program AC	-0.0325	0.0220	0.8446	0.0805	-0.0012	-0.7490	-1.5069	-0.0762
	-0.9193	0.0554	0.6069	0.4437	-0.0025	-1.1550	-1.0524	
Program AD	0.0126	0.0536	0.0055	0.0062	-0.1212	0.1880	0.0241	-0.2813
	1.0321	0.3913	0.0114	0.0988	-0.7526	0.8409	0.0487	

Table 3: Regression results for the Multi-Factor model using the Volatility Proxies

OEX is the S&P 100, NOTE is the futures contract on the 5-year Treasury note, GSCI is the futures on the GSCI contract, and EURO is the futures on the Euro-currency.

Dependent Variable	ALPHA	OEX	NOTE	GSCI	EURO	Adj. R-squared
Program A	-0.0018	-0.0075	-0.0061	0.0167	0.0440	0.102125
	-0.0600	-0.2219	-0.6527	1.4867	2.1088	
Program B	-0.0116	0.0316	0.0041	0.0018	0.0178	0.3393091
	-0.9348	2.1987	1.0406	0.3724	2.0158	
Program C	-0.0266	0.0129	0.0098	0.0113	0.0227	0.3954676
	-1.8019	0.7535	2.0677	1.9943	2.1552	
Program D	-0.0196	0.0302	0.0074	0.0132	0.0240	0.3087007
	-1.0580	1.4029	1.2498	1.8477	1.8099	
Program E	-0.0024	0.0033	0.0005	-0.0011	0.0002	-0.1776552
	-0.3480	0.4153	0.2170	-0.4119	0.0449	
Program F	-0.0004	0.0006	-0.0040	-0.0014	0.0129	0.3914076
	-0.0856	0.1038	-2.5646	-0.7520	3.7126	
Program G	-0.0175	-0.0217	-0.0016	0.0042	-0.0151	0.091409
	-1.1735	-1.2521	-0.3441	0.7344	-1.4216	
Program H	-0.0026	-0.0056	0.0021	-0.0015	-0.0037	-0.0289677
	-0.5312	-0.9709	1.2963	-0.7660	-1.0430	
Program I	-0.0062	0.0034	0.0017	-0.0026	-0.0038	-0.04644
	-1.0545	0.4989	0.8755	-1.1477	-0.9018	
Program J	-0.0131	0.0122	0.0021	0.0073	0.0212	0.2968242
	-1.1624	0.9292	0.5836	1.6901	2.6181	
Program K	-0.0592	-0.0058	0.0031	0.0302	0.0455	0.171587
	-1.6568	-0.1387	0.2659	2.2021	1.7803	
Program L	-0.0148	-0.0101	0.0006	-0.0019	0.0074	0.2674721
	-3.9501	-2.3336	0.4906	-1.2904	2.7559	
Program M	-0.0147	0.0072	-0.0012	0.0014	-0.0018	-0.1114596
	-2.6077	1.1000	-0.6803	0.6307	-0.4375	
Program N	-0.0035	0.0211	-0.0017	0.0051	0.0094	-0.0371978
	-0.2307	1.2100	-0.3431	0.8872	0.8796	
Program O	-0.0049	0.0066	-0.0032	-0.0026	0.0421	0.7191861
	-0.6316	0.7302	-1.2599	-0.8543	7.5460	
Program P	-0.0127	0.0009	0.0040	-0.0024	0.0092	-0.0319193
	-1.0556	0.0619	1.0303	-0.5193	1.0734	
Program Q	-0.0134	-0.0038	0.0044	-0.0023	0.0072	-0.0851865
	-1.0143	-0.2479	1.0469	-0.4482	0.7671	
Program R	0.0095	0.0849	-0.0089	0.0132	0.0662	0.25085
	0.2484	1.9020	-0.7266	0.8954	2.4155	
Program S	0.0533	-0.0120	-0.0083	-0.0190	0.0209	-0.0246177
	1.5350	-0.2975	-0.7415	-1.4246	0.8401	
Program T	-0.0172	0.0135	0.0093	0.0038	0.0152	0.1959737
	-1.0931	0.7396	1.8457	0.6337	1.3509	
Program U	-0.0028	0.0012	-0.0020	-0.0009	-0.0026	-0.1207321

Dependent Variable	ALPHA	OEX	NOTE	GSCI	EURO	Adj. R-squared
	-0.4014	0.1526	-0.8875	-0.3357	-0.5229	
Program V	-0.0140	-0.0044	0.0034	-0.0065	0.0083	-0.0587078
	-0.9866	-0.2687	0.7388	-1.1928	0.8149	
Program W	-0.0025	0.0049	-0.0017	-0.0008	-0.0043	-0.1425878
	-0.2844	0.4700	-0.5866	-0.2337	-0.6791	
Program X	0.0710	0.0689	-0.0168	0.0183	0.0409	-0.0568205
	1.1910	0.9938	-0.8804	0.7989	0.9603	
Program Y	-0.0494	0.0105	0.0124	0.0226	0.0208	0.5864347
	-3.4467	0.6277	2.6974	4.1065	2.0365	
Program Z	0.1045	0.0813	-0.0194	-0.0275	0.0374	0.2861271
	2.8265	1.8911	-1.6316	-1.9368	1.4156	
Program AA	-0.0396	0.0167	0.0072	0.0164	0.0189	0.387522
	-2.6459	0.9580	1.5015	2.8603	1.7652	
Program AB	-0.0010	0.0174	-0.0005	0.0094	0.0133	-0.0195961
	-0.0558	0.8031	-0.0799	1.3139	0.9956	
Program AC	0.0058	0.0122	-0.0030	-0.0029	0.0005	-0.1828502
	0.2381	0.4310	-0.3851	-0.3119	0.0310	
Program AD	0.0163	0.0122	0.0010	-0.0009	-0.0123	0.181672
	2.5433	1.6350	0.4898	-0.3735	-2.6883	

Table 4: Regression Results for Multi-factor Model using Market Sector Benchmarks and Volatility Proxies

ALPHA is the excess return, TRI is the S&P 500 Total Return Index, LUA is the Lehman U.S. Aggregate, GSCIE is the GSCI Energy, GSCIM is the GSCI Metals, GSCIN is the GSCI Non-energy, NEXN is the JP Morgan Nominal Exchange Rate Index, OEX is the S&P 100, NOTE is the futures contract on the 5-year Treasury note, GSCI is the futures on the GSCI contract, and EURO is the futures on the Euro-currency.

Dependent Variable	Intercept	TRI	LUA	GSCIE	GSCIM	GSCIN	NEXN	OEX	NOTE	GSCI	EURO	Adj. R-squared
Program A	-0.0029	-0.8105	1.1514	-0.0780	0.9850	0.0908	0.2699	-0.0208	-0.0104	0.0216	0.0359	-0.0846
	-0.0510	-1.2278	0.5262	-0.2555	1.3806	0.0909	0.1210	-0.4924	-0.8556	1.3171	1.2982	
Program B	-0.0217	-0.2343	-0.2208	0.1697	0.2216	0.2301	-0.5977	0.0237	0.0017	-0.0021	0.0160	0.4298
	-1.0665	-0.9916	-0.2819	1.5527	0.8676	0.6436	-0.7488	1.5668	0.4007	-0.3633	1.6129	
Program C	-0.0294	-0.1223	-0.1247	0.2397	-0.2585	0.9357	-0.7434	0.0137	0.0048	0.0079	0.0218	0.5920
	-1.3719	-0.4909	-0.1510	2.0796	-0.9600	2.4819	-0.8831	0.8571	1.0364	1.2802	2.0931	
Program D	-0.0357	0.0721	0.4519	0.0269	0.1594	0.1808	-1.1729	0.0236	0.0088	0.0113	0.0176	0.1205
	-0.9650	0.1676	0.3169	0.1351	0.3428	0.2778	-0.8072	0.8596	1.1154	1.0591	0.9771	
Program E	-0.0196	-0.0046	-0.0762	-0.0701	0.0297	-0.0836	-0.8142	-0.0006	0.0014	-0.0006	-0.0020	-0.3239
	-1.5152	-0.0305	-0.1528	-1.0085	0.1830	-0.3676	-1.6028	-0.0619	0.5124	-0.1650	-0.3192	
Program F	-0.0019	-0.2231	0.2975	-0.0548	0.2180	0.0319	0.0445	-0.0021	-0.0053	0.0009	0.0110	0.4116
	-0.2192	-2.2609	0.9095	-1.2010	2.0439	0.2134	0.1335	-0.3286	-2.9042	0.3515	2.6527	
Program G	-0.0025	0.6903	-0.6596	0.0437	-0.2507	0.4672	0.7277	-0.0071	0.0038	0.0041	-0.0055	0.3438
	-0.1137	2.6483	-0.7634	0.3628	-0.8898	1.1846	0.8265	-0.4282	0.7909	0.6391	-0.4996	
Program H	0.0042	0.0670	-0.0589	0.0023	-0.0748	0.0734	0.3522	-0.0020	0.0022	-0.0009	-0.0017	-0.3697
	0.4136	0.5689	-0.1510	0.0424	-0.5877	0.4120	0.8855	-0.2660	1.0146	-0.3090	-0.3433	

Dependent Variable	Intercept	TRI	LUA	GSCIE	GSCIM	GSCIN	NEXN	OEX	NOTE	GSCI	EURO	Adj. R-squared
Program I	-0.0003	0.0434	0.3751	0.0073	0.0698	-0.0496	0.1699	0.0031	0.0023	-0.0029	-0.0056	-0.3907
	-0.0269	0.3102	0.8086	0.1127	0.4612	-0.2342	0.3594	0.3514	0.8859	-0.8372	-0.9538	
Program J	-0.0106	-0.4756	0.7630	0.1305	0.1302	0.4841	-0.3155	0.0065	-0.0041	0.0066	0.0146	0.6419
	-0.7423	-2.8667	1.3874	1.7009	0.7260	1.9284	-0.5628	0.6138	-1.3574	1.5912	2.0948	
Program K	-0.0365	-0.7763	0.1521	0.4117	-0.1253	0.1703	0.6819	-0.0125	-0.0089	0.0226	0.0428	0.0927
	-0.5511	-1.0089	0.0596	1.1569	-0.1507	0.1463	0.2623	-0.2535	-0.6308	1.1835	1.3298	
Program L	-0.0197	0.0445	0.0662	0.0220	-0.0880	0.1411	-0.4428	-0.0106	0.0004	-0.0026	0.0058	0.3336
	-3.1246	0.6077	0.2723	0.6500	-1.1111	1.2721	-1.7882	-2.2686	0.3168	-1.4236	1.9008	
Program M	-0.0180	-0.0419	-0.5037	0.0608	-0.0900	0.1346	-0.0917	0.0077	-0.0024	0.0003	0.0009	-0.2619
	-1.6957	-0.3398	-1.2336	1.0669	-0.6761	0.7223	-0.2205	0.9754	-1.0575	0.0873	0.1821	
Program N	-0.0349	-0.1861	1.4159	-0.0854	0.5758	-0.0824	-2.0285	0.0032	-0.0001	0.0047	-0.0068	0.2621
	-1.5592	-0.7140	1.6392	-0.7083	2.0442	-0.2089	-2.3043	0.1947	-0.0292	0.7291	-0.6247	
Program O	0.0046	-0.1610	0.7638	-0.0480	0.4472	-0.3150	0.5887	0.0020	-0.0025	-0.0014	0.0382	0.7909
	0.3881	-1.1632	1.6647	-0.7490	2.9892	-1.5040	1.2590	0.2213	-0.9821	-0.3999	6.5905	
Program P	0.0163	-0.1455	0.0430	0.0847	-0.3018	0.6112	1.3205	0.0126	-0.0010	0.0000	0.0145	0.0111
	0.7838	-0.5999	0.0535	0.7554	-1.1514	1.6657	1.6117	0.8126	-0.2311	0.0069	1.4292	
Program Q	0.0203	-0.1267	0.1827	0.0983	-0.3660	0.7686	1.4337	0.0101	-0.0011	0.0006	0.0125	-0.0054
	0.9037	-0.4845	0.2107	0.8128	-1.2946	1.9423	1.6228	0.6031	-0.2318	0.0854	1.1411	
Program R	0.0166	-1.5728	1.3273	0.2437	0.7685	1.1197	0.0797	0.0691	-0.0260	0.0154	0.0553	0.3399
	0.2606	-2.1232	0.5406	0.7112	0.9599	0.9991	0.0318	1.4594	-1.9084	0.8351	1.7817	
Program S	0.0836	0.4247	-1.4996	-0.0447	0.9133	0.1635	3.0001	0.0009	0.0007	-0.0141	0.0421	0.2695
	1.6104	0.7037	-0.7498	-0.1602	1.4004	0.1791	1.4718	0.0241	0.0630	-0.9398	1.6643	
Program T	-0.0115	0.0588	-0.2950	0.3167	-0.3252	0.6648	-0.3485	0.0152	0.0059	-0.0027	0.0161	0.2600
	-0.4308	0.1895	-0.2866	2.2049	-0.9690	1.4148	-0.3321	0.7653	1.0278	-0.3506	1.2369	
Program U	-0.0080	0.2674	-0.3875	0.0735	-0.2640	0.3524	-0.5034	0.0049	-0.0010	-0.0028	-0.0006	0.1691
	-0.7466	2.1560	-0.9426	1.2811	-1.9697	1.8782	-1.2016	0.6152	-0.4166	-0.9174	-0.1121	
Program V	-0.0236	0.2462	-1.1121	0.2610	-0.6783	0.1165	-0.9123	-0.0032	0.0019	-0.0150	0.0117	0.0952
	-1.0184	0.9136	-1.2447	2.0938	-2.3283	0.2858	-1.0019	-0.1880	0.3841	-2.2378	1.0329	
Program W	0.0064	0.1587	0.4240	0.2421	-0.3525	0.1900	-0.3927	0.0045	-0.0034	-0.0076	-0.0086	0.3670
	0.5408	1.1597	0.9347	3.8242	-2.3833	0.9178	-0.8495	0.5131	-1.3456	-2.2409	-1.4979	
Program X	0.0216	-0.8461	-1.0470	0.2989	1.5667	1.5846	-1.8930	0.0450	-0.0210	0.0163	0.0403	0.0622
	0.2174	-0.7327	-0.2735	0.5597	1.2552	0.9070	-0.4852	0.6092	-0.9866	0.5670	0.8324	
Program Y	-0.0434	0.2915	-0.8790	0.1589	-0.6346	0.7520	-0.0196	0.0223	0.0108	0.0206	0.0284	0.5758
	-1.6928	0.9770	-0.8889	1.1514	-1.9680	1.6659	-0.0195	1.1696	1.9710	2.7795	2.2745	
Program Z	0.1161	0.8962	-1.0862	0.8948	-1.6939	0.7058	-1.6344	0.0873	-0.0229	-0.0536	0.0384	0.4256
	1.9790	1.3136	-0.4804	2.8359	-2.2974	0.6838	-0.7092	2.0008	-1.8235	-3.1621	1.3426	
Program AA	-0.0447	-0.4532	-1.4930	-0.1119	-0.3163	0.2608	0.5907	0.0245	0.0013	0.0227	0.0303	0.5561
	-1.9855	-1.7307	-1.7203	-0.9240	-1.1177	0.6584	0.6679	1.4616	0.2769	3.4875	2.7616	
Program AB	-0.0093	-0.3325	2.0936	0.1052	0.3415	0.8378	-1.5645	0.0061	-0.0047	0.0086	-0.0044	0.4230
	-0.3773	-1.1540	2.1920	0.7893	1.0965	1.9217	-1.6073	0.3310	-0.8826	1.2022	-0.3647	
Program AC	-0.0222	0.0476	1.3926	0.2751	0.0288	-1.1306	-2.6841	-0.0156	-0.0018	-0.0191	-0.0220	-0.0723
	-0.5397	0.0998	0.8801	1.2457	0.0558	-1.5653	-1.6644	-0.5111	-0.2089	-1.6069	-1.0997	
Program AD	0.0219	0.1120	0.4396	0.0182	-0.1309	0.1845	-0.0629	0.0146	0.0010	-0.0010	-0.0144	0.0007
	1.7527	0.7691	0.9109	0.2703	-0.8319	0.8376	-0.1279	1.5724	0.3751	-0.2899	-2.3671	

We took the results for the risk and performance measures, sorted by time frame and each measure was averaged. The results are shown in Tables 5—8 below.

Table 5: Average Values for Different Performance and Risk Measures across Time Frames

ST-MT is the difference between the short-term and the intermediate-term time frame. MT-LT is the difference between the intermediate-term and the long-term time frame. ST-LT is the difference between the short-term and the long-term time.

				AVERAGE			
Time Frame	Mean ROR	Standard Deviation	Downside Deviation	Maximum Drawdown	Sharpe Ratio	Sortino Ratio	Calmar Ratio
Short term	1.90%	5.74%	2.48%	-34.54%	-0.0251	2.0585	0.0834
Medium term	1.62%	6.37%	2.75%	-19.93%	-0.0553	0.7380	0.1448
Long term	0.78%	6.20%	2.71%	-39.60%	-0.2915	0.1930	0.0293
ST-MT	0.28%	-0.62%	-0.26%	-14.60%	0.0302	1.3204	-0.0614
MT-LT	0.84%	0.17%	0.04%	19.67%	0.2362	0.5451	0.1156
ST-LT	1.11%	-0.45%	-0.23%	5.06%	0.2664	1.8655	0.0542

Table 6: p-values for t-tests for Different Performance and Risk Measures.

ST/MT is the p-value between the short-term and the intermediate-term time frame. MT-LT is the p-value between the intermediate-term and the long-term time frame. ST-LT is the p-value between the short-term and the long-term time frame.

t-tests	Mean ROR	Standard Deviation	Downside Deviation	Maximum Drawdown	Sharpe Ratio	Sortino Ratio	Calmar Ratio
ST/MT	0.3748	0.4966	0.2813	0.3734	0.3455	0.0785	0.3946
MT/LT	0.0832	0.4620	0.3981	0.4689	0.4811	0.0479	0.0735
ST/LT	0.0906	0.4700	0.2597	0.4224	0.3957	0.3506	0.0613

For Mean ROR in Table 5 we notice that ST-MT, MT-LT, ST-LT are all positive indicating that the shorter-term programs in our sample perform better on average than the longer-term programs. When we look at the p-values in Table 6 we see that they are 0.3748, 0.0832, and 0.0906 respectively. This indicates that there is no statistical difference between short-term and medium-term, but there is a statistical difference between intermediate term and short-term, and short-term and long-term. However, the p-values for standard deviation, downside deviation, and maximum drawdown do not reflect any difference in the performance across time frame. For the Sortino and Calmar ratios there is evidence of a difference in performance for different time frames. However the results are mixed.

Table 7: Average Values of Excess Return for Different Multi-factor Models across Time Frames

Alpha 1 is the average excess return for the market sector benchmark multifactor model. Alpha 2 is the average excess return for the volatility proxies multifactor model. Alpha 3 is the average excess return for the market sector benchmarks and volatility proxies multifactor model.

	AVERAGE		
Time Frame	Alpha 1	Alpha 2	Alpha 3
Short term	-0.0074	-0.0036	-0.0055
Medium term	-0.0036	0.0095	0.0085
Long term	-0.0199	-0.0193	-0.0146
ST-MT	-0.0038	-0.0132	-0.0140
MT-LT	0.0163	0.0288	0.0231
ST-LT	0.0125	0.0157	0.0091

Table 8: p-values for t-tests for Multi-Factor Models

t-tests	Alpha 1	Alpha 2	Alpha 3
ST/MT	0.3619	0.1847	0.1960
MT/LT	0.0782	0.0234	0.0897
ST/LT	0.0487	0.0971	0.1937

Our null hypothesis is that the excess return for the short-term, intermediate-term, and long-term time frame programs are the same. Looking at the Table 7 we see that the mean difference in Alpha 1 for ST-MT is –0.0038, indicating that on average the intermediate-term programs perform better than the short-term programs. However, the corresponding p-value in Table 8 is high at 0.3619, meaning that this not significant. For MT-LT and ST-LT the values of the difference in Alpha 1 are 0.0163 and 0.0125 respectively. This indicates that the intermediate-term programs do better than the long-term programs, and the short-term programs do better than the long-term programs on average based on the first multi-factor model. The p-values are 0.0782 and 0.0487 showing that there is a very small possibility of making a Type I error. The results are similar for the multi-factor model involving volatility proxies. Thus we are able to reject the null hypothesis that the alphas of the multi-factor models are the same with 90-95% confidence for MT-LT and ST-LT.

5 Conclusions and Directions for Further Research

We were able to reject the null hypothesis that short-term price movements may be due to only noise and CTAs have equal performance regardless of time frame. We reject this hypothesis based on the significance of the p-values for the unadjusted returns and the Calmar ratios, as well as from the α's from the multi-factor models.

For the future we would like to extend this research as follows. We would like to

- increase sample size of CTA programs,
- look at performance measures over different time periods, and
- investigate the correlation of these groups to traditional assets (stocks and bonds) to examine their relative ability to reduce portfolio risk through diversification.

Bibliography

About the Commodity Futures Trading Commission. 22 April 2004. Commodity Futures Trading Commission. 24 April 2004. <http://www.cftc.gov/cftc/cftcabout.htm>

Arsad, Zainudin and J. Andrew Coutts. The Weekend Effect, Good News, Bad News and the Financial Times Industrial Ordinary Shares Index: 1935–94. *Applied Economics Letters*, 3:797-801, 1996, 3.

Brinson, Gary P., L. Randolph Hood, and Gilbert L. Beebower. Determinants of Portfolio Performance. *Financial Analysts Journal*, July/August 1986.

Brinson, Gary P., Brian D. Singer, and Gilbert L. Beebower. Determinants of Portfolio Performance II. *Financial Analysts Journal*, pages 40-48, May/June 1991.

Burghardt, Galen, Ryan Duncan, and Lianyan Liu. Understanding Drawdowns. Technical report, Carr Futures, May 2003.

Burghardt, Galen and Lianyan Liu. How important are daily return data? Technical report, Carr Futures, April 2003.

Diz, Fernando. CTA Survivor and Nonsurvivor: An Analysis of Relative Performance. *The Journal of Alternative Investments*, pages 57-71, Summer 1999.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake. Survivorship Bias and Mutual Fund Performance. *The Review of Financial Studies*, 9(4): 1097-1120, 1996.

Fung, William and David A. Hsieh. Survivorship bias and investment style in the returns of CTAs. *The Journal of Portfolio Management*, 24(1): 30-41, Fall 1997.

Fung, William and David A. Hsieh. Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases. *Journal of Financial and Quantitative Analysis*, 35(3): 291-307, September 2000.

Ghaleb-Harter, Tanya. The Case for Global Macro. DB Absolute Return Strategies Research, 2004(?).

Harding, David, Georgia Nakou, and Ali Nejjar. The Pros and Cons of

"Drawdown" as a Statistical Measure of Risk for Investments. Winton Capital Management, 2002.

Jaeger, Robert A. Risk: Defining it, Measuring it, and Managing it. In *Managing Hedge Fund Risk*. Risk Books, 2000.

Kazemi, Hossein and Thomas Schneeweis. Conditional Performance of Hedge Funds. University of Massachusetts, March 2003.

Kritzman, Mark. What practitioners need to know . . . about estimating volatility part 1. *Financial Analysts Journal*, 47(4):22–25, July–August 1991.

Kritzman, Mark. What practitioners need to know . . . about estimating volatility part 2. *Financial Analysts Journal*, 47(5):10–11, September–October 1991.

Kritzman, Mark. What practitioners need to know . . . about regressions. *Financial Analysts Journal*, 47(3):12–15, May–June 1991.

Kritzman, Mark. What practitioners need to know . . . about factor methods. *Financial Analysts Journal*, 49(1):12–15, January–February 1993.

Kritzman, Mark. Towards defining an asset class. *The Journal of Alternative Investments*, pages 79-82, Summer 1999.

Lungarella, Gildo – Harcourt AG. Strategy Focus… Managed Futures: A Real Alternative. *swissHEDGE*, 4th Quarter 2002.

Lungarella, Gildo – Harcourt AG. Strategy Focus… Investing in Global Macro. *swissHEDGE*, 1st Quarter 2004.

Malek, Marc and Sergei Dobrovolsky. Market volatility and CTA performance. Technical report, Camelot Partners, February 2003.

Nakou, Georgia and Duncan Brown – Winton Capital Management. Investors View… Beauty or the Beast? Myth and Reality about Managed Futures. *swissHEDGE*, 4th Quarter 2002.

Rulle, Michael S. Trend following: Performance, Risk and Correlation Characteristics. MFA Reporter, pages 1–2,10–12, February/March 2003.

Schneeweis, Thomas and Richard Spurgin. Survivor Bias in Managed Futures Research.

Schneeweis, Thomas, Richard Spurgin, and Mark Potter. Managed futures and hedge fund investment for downside equity risk management. *Derivatives*

Quarterly, pages 1–11, Fall 1996.

Sharpe, William F. Integrated asset allocation. *Financial Analysts Journal*, pages 25–32, September–October 1987.

Sharpe, William F. Asset Allocation: Management Style and Performance Measurement. *The Journal of Portfolio Management*, pages 7-19, Winter 1992.

Sortino, Frank A. and Hal J. Forsey. On the use and misuse of downside risk. *The Journal of Portfolio Management*, pages 35–42, Winter 1996.

Sortino, Frank A. and Lee N. Price. Performance measurement in downside risk framework. *The Journal of Investing*, pages 59–64, Fall 1994.

Sortino, Frank A. and Robert van der Meer. Downside risk. *The Journal of Portfolio Management*, pages 27–31, Summer 1991.

Spurgin, Richard. A benchmark for commodity trading advisor performance. Technical report, CISDM Working Paper Series, 1999.

Spurgin, Richard. A Benchmark for Commodity Trading Advisor Performance. *The Journal of Alternative Investments*, pages 11-21, Summer 1999.

VIX: CBOE Volatility Index. 18 September 2003. Chicago Board Options Exchange. 27 April 2004. <http://www.cboe.com/micro/vix/vixwhite.pdf>

Wilkens, Kathryn. Evidence of Risk/Returns Patterns in Cash and Futures Markets. *The Journal of Alternative Investments*, pages 45-67, Fall 2000.

Appendix A

Below is the entire survey that was sent to the managed futures managers in the ITR CTA database.

CTA Performance Evaluation

I am a graduate student in the Department of Mathematical Sciences at Worcester Polytechnic Institute working on my M.S. degree in Financial Mathematics. For my thesis I am researching the impact of time frame and related style factors on CTA performance. If you participate, in exchange for your input, I will provide you with a final copy of my thesis. All responses will be kept strictly confidential. I expect this research to shed some light on the role time frame plays in the relative performance of systematic traders. I look forward to sharing the results of my research with the managers who participate, so I hope you will be among them. Thank you very much. If you have any questions or need further clarification please contact me at nthomas@wpi.edu.

Please provide us with the following information

Company:

Program:

Name:

E-mail:

1. Is your program:

Trend following	Percentage	%
Non-trend following	Percentage	%
	Please specify:	
Counter trend	Percentage	%

2. What is the composition of your decision-making process?

Systematic %

Discretionary %

Systematic and discretionary Comment:

3. What market sectors do you trade?

Equities	Percent of trading	%
Fixed Income	Percent of trading	%
Foreign Exchange	Percent of trading	%
Commodities	Percent of trading	%

4. What is the frequency of the data you use to generate trading signals? (e.g. 5-min, daily, weekly)

5. What is your average holding period?

Overall

For winning trades

For losing trades

6. For winning trades, in what proportions on average over the last 12 months have you employed the following time frames?

0 - 2 days	Percent of trading	%
3 - 10 days	Percent of trading	%
11 - 30 days	Percent of trading	%
31 - 80 days	Percent of trading	%
80 days or more	Percent of trading	%

7a. Has your overall mix of time frames changed, resulting in a different average holding period? <Select One>

7b. If so, when did they change?

8. Approximately what percentage change in price does it take on average to trigger your entry into the market to trade (in terms of the average holding period for each model, where relevant)?

0 - 2 days	Percent change	%
3 - 10 days	Percent change	%
11 - 30 days	Percent change	%
31 - 80 days	Percent change	%
80 days or more	Percent change	%

9. If your models are driven by volatility, what percentage change in volatility does it take on average to trigger your entry into the market to trade (in terms of the average holding period for each model, where relevant)?

0 - 2 days	Percent change	%
3 - 10 days	Percent change	%
11 - 30 days	Percent change	%
31 - 80 days	Percent change	%
80 days or more	Percent change	%

10. What is the average percentage risk per trade in account/fund equity that you take (in terms of average holding period of the various models, where relevant)?

0 - 2 days	Average risk	
3 - 10 days	Average risk	
11 - 30 days	Average risk	
31 - 80 days	Average risk	
80 days or more	Average risk	

11a. Do you always reverse positions or do you have a neutral mode? <Select One>

11b. If you have a neutral mode in your model(s), approximately what percentage of the time are you in the markets? %

12. If you require any clarification about any of the questions please let me know. Please provide any additional information/comments about your program here as well as any comments you have about the survey: